

# Перспективы развития открытых проектов в области машинного обучения

Шмырёв Н.В.

АЦ Технологии

# Библиотека распознавания речи Vosk

<https://github.com/alphacep/vosk-api>

- Несколько платформ - Linux, Win, OSX, Android, iOS, RPi
- Поддерживает 20+ языков и диалектов. Последние - **хинди, японский**
- Несколько языков программирования и сред - Python, Node, C#, C, Go, Java, **Unity, Flutter**
- Системы телефонии - Asterisk, Freeswitch, Jitsi
- Модели разного размера для точного распознавания на серверах и мобильных устройствах

## Библиотека распознавания речи Воск

- **Эффективное декодирование на GPU**
- Поточковая обработка с минимальным временем отклика
- Настраиваемый словарь для распознавания специальных терминов
- Идентификация говорящего
- Пунктуация
- **Ассистент и системная клавиатура на AOSP (LineageOS, ...)**

## Эволюция проектов машинного обучения

- академические проекты (2000-2010)
- академические проекты в сотрудничестве с индустрией (2010-2017)
- крупные индустриальные проекты (2018-2022)
- проекты-публикации (2020-2022).

# Синтез реальности

<http://soulmachines.com>

**Create your Digital Workforce**  
Customizable with Smart Skills

**LEARN MORE**

The banner features a central cluster of ten hexagonal icons, each containing a different digital skill icon and label: 'Big Data', 'Time', 'Messaging', 'Integration', 'Personalization', 'Automation', 'Shopping', 'Communication', 'Analytics', and 'Real-time'. The background is a blurred image of a woman in a white top, looking towards the camera.

## Высокая стоимость тренировки



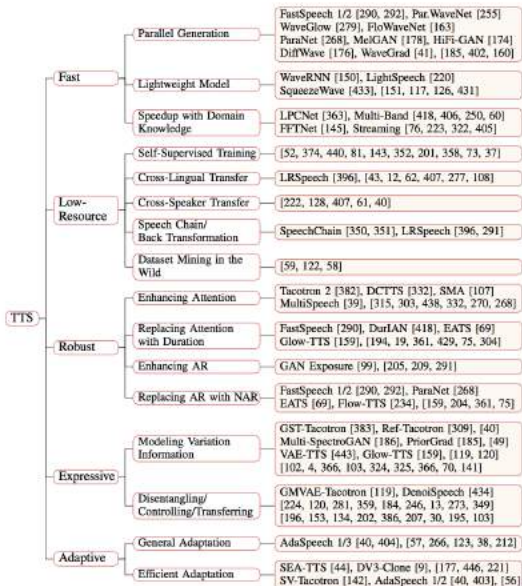
## Взрывной рост проектов

- 20+ проектов распознавания речи
- 40+ проектов синтеза
- и так далее для каждой области

Facebook Wav2Vec, Facebook/Flashlight, Facebook/Fairseq,  
Google/Trax, DeepSpeech, Kaldi, Xiaomi/K2, Bytedance/NeurST,  
Baidu/PaddlePaddle, Nvidia/Nemo, CMU/Espnet, Tencent/Pika,  
RWTH/Rasr, Didi/Athena....

# Survey on Neural Speech Synthesis

<https://arxiv.org/pdf/2106.15561.pdf>

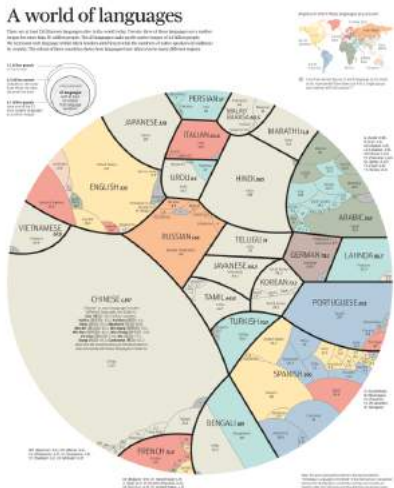




## Перспективы

- Задачи всемирного масштаба
- Проекты-интеграции
- Специализация в предметных областях

# Перспективы. Задачи всемирного масштаба



## Перспективы. Интеграции

Huggingface, Python Speech Recognition, Nvidia Nemo, FastAPI



## Перспективы. Специализация



## Перспективы

- Задачи всемирного масштаба
- Проекты-интеграции
- Специализация в предметных областях

## Вопросы/Контакты

Сайт: <https://alphacephei.com>

Код: <https://github.com/alphacep>

Телеграмм: [https://t.me/speech\\_recognition](https://t.me/speech_recognition)

Почта: [contact@alphacephei.com](mailto:contact@alphacephei.com)